APPROXIMATIONS TO THE AVERAGE RUN LENGTHS OF CUSUM TESTS

by

Moshe Pollak
Hebrew Univeristy

and

David Siegmund
Stanford University

TECHNICAL REPORT NO. 37
AUGUST 1986

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

DTIC
ELECTE
OCT 7 1986
B

86 10 7 060

APPROXIMATIONS TO THE AVERAGE RUN LENGTHS
OF CUSUM TESTS

by

Moshe Pollak
Hebrew University

and

David Siegmund
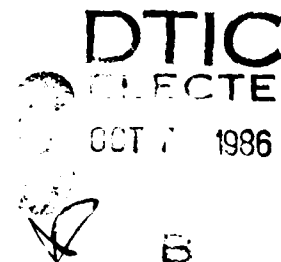Stanford University

TECHNICAL REPORT NO. 37
AUGUST 1986

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

NTIS
DTIC
Unan
J

By
Dist

A-1

APPROXIMATIONS TO THE AVERAGE RUN LENGTHS OF CUSUM TESTS

by

Moshe Pollak

Hebrew University

and

David Siegmund

Stanford University

Abstract

Simple approximations are given for the average run lengths of CUSUM tests: both with and without the fast initial response feature, and starting from a quasi-stationary state. Numerical examples illustrate the accuracy of the results.

1. Introduction.

For $\nu = 0, 1, 2, \cdots, = \infty$ let P_ν denote the probability measure under which $y_1, y_2, \cdots, y_\nu$ are independently distributed with probability density function g_0 and $y_{\nu+1}, \cdots$ are independently distributed according to g, which is assumed to belong to some family of probability density functions. Let g_1 denote one particular possibility from the set of densities g, let $x_k = \log[g_1(y_k)/g_0(y_k)]$, and $S_n = x_1 + \cdots + x_n$. A CUSUM test for detecting the change-point ν is defined by a stopping rule of the form

$$\tau = \inf\{n : S_n - \min_{k \le n} S_k \ge b\}.$$

See vanDobben de Bruyn (1968) for a background discussion of CUSUM tests.

The stopping rule τ is traditionally evaluated in terms of its average run lengths: $E_\infty(\tau)$, the in control average run length, which should be large, and $E_0(\tau)$, the out of control average run length, which should be small. Pollak and Siegmund (1985) suggest that one also consider $E_\nu(\tau - \nu \mid \tau > \nu)$ for large ν, i.e. the average run length starting from a quasi-stationary state.

Assuming that g_0 and g can be imbedded in an exponential family of distributions, Siegmund (1985, Theorem 10.16) has given approximations for $E_\infty(\tau)$ and $E_0(\tau)$. The main result

of this note is an approximation for $E_\nu(\tau - \nu \mid \tau > \nu)$ which is valid when ν and b are large. We also improve Siegmund's (1985) approximations.

To provide a "fast initial response" to detect changes occurring at $\nu = 0$, Lucas and Crosier (1982) modify the definition of τ by putting $S_0 = x$, $S_n = x + x_1 + \cdots + x_n$, and

$$\tau = \inf\{n : S_n - \min_{\sigma_0 \le k \le n} S_k \ge b\},$$

where $\sigma_0 = \inf\{n : S_n \le 0\}$. We also give approximations for the average run lengths of this modification and show by a numerical example how our approximations can provide guidelines for selecting the value of x.

A more precise description of the approximations is given in Section 2. Section 3 contains numerical examples to illustrate the accuracy of the approximations.

Pollak and Siegmund (1985) in the context of Brownian motion studied the Shiryayev-Roberts procedure as a possible competitor to CUSUM tests. It would be desirable to have approximations similar to those of the present paper for that process. Pollak's (1987) approximation for the P_∞ average run length is adequate, but otherwise it seems a difficult problem to develop simple, comparably good approximations for the Shiryayev-Roberts process.

2. Approximations

It will be convenient to give our results in a canonical form, which usually is not the most suitable form for any particular application, except perhaps for tests for a change in a normal mean. We develop here our notation and definitions ab *initio*. They are not necessarily consistent with Section 1.

For θ in some real interval containing 0, let F_θ denote a probability distribution of the form

$$dF_\theta(x) = \exp[\theta x - \psi(\theta)]dF_0(x).$$

Then $\psi'(\theta) = \int x dF_\theta$ and $\psi''(\theta) = \int x^2 dF_\theta - [\psi'(\theta)]^2 > 0$. By an affine transformation if necessary one may assume without loss of generality that $\psi'(0) = 0$ and $\psi''(0) = 1$. Then

$$sgn(\theta) = sgn\psi'(\theta).$$

2

For each $\theta_0 < 0$ $(\theta_1 > 0)$ assume there exists $\theta_0^* > 0$ $(\theta_1^* < 0)$ such that $\psi(\theta_i) = \psi(\theta_i^*)$ $(i = 0, 1)$. By the convexity of ψ the value of θ_0^* (θ_1^*) is unique. Let $\Delta_i = \theta_i - \theta_i^*$. Note that $\Delta_0 < 0 < \Delta_1$.

Let $\theta_0 < 0 \leq \theta_1$ and $\nu = 0, 1, \cdots, +\infty$. Let P_ν denote the probability under which $x_1, \cdots, x_\nu$ are independent with distribution F_{θ_0} and $x_{\nu+1}, \cdots$ are independent with distribution F_{θ_1}. Of course P_ν depends also on θ_0 and θ_1, but we suppress this dependence to simplify the notation. When $\nu = 0$ or ∞ so the entire sequence $x_1, x_2, \cdots$ is independent and identically distributed, it will be convenient to write P_θ to emphasize dependence on θ. (Note that according to this convention P_0 now denotes P_θ for $\theta = 0$, *not* P_ν for $\nu = 0$.)

Let $0 \leq x < b$, $S_n = x + x_1 + \cdots + x_n$, and

$$\tau = \inf\{: S_n - \min_{\sigma_0 \leq k \leq n} S_k \geq b\},$$

where $\sigma_0 = \inf\{n : n \geq 0, S_n \leq 0\}$. When $x > 0$ we shall write P_θ^x to denote dependence of probabilities on the value of x.

It will be convenient to use the following additional notation. Let

$$\tau_+(\tau_-) = \inf\{n : S_n > 0(S_n < 0)\}.$$

Put $\rho_1 = E_0(S_{\tau_+}^2)/[2E_0(S_{\tau_+})]$ and $\rho_0 = E_0(S_{\tau_-}^2)/[2E_0(S_{\tau_-})]$. In general ρ_0 and ρ_1 must be computed numerically. For an algorithm see Siegmund (1985, Theorem 10.55 and Problem 10.7). In the case of a standard normal distribution $\rho_1 = -\rho_0 \cong .583$. Also let $\gamma_1 = E_0(S_{\tau_+}^3)/[3E_0(S_{\tau_+})] - \rho_1^2$ and $\gamma_2 = E_0(S_{\tau_-}^3)/[3E_0(S_{\tau_-})] - \rho_0^2$. Like ρ_0 and ρ_1 the constants γ_0 and γ_1 must in general be computed numerically. However, their role is comparatively minor and one might choose to neglect them. For continuous and symmetric F_0, hence in the normal case, it is easy to use the Wiener-Hopf factorization of the distributions of S_{τ_+} and S_{τ_-} (e.g. Siegmund, 1985, Theorem 8.41) to show that $\gamma_1 = \gamma_0 = E_0(x_1^4)/12$.

We shall assume throughout that F_0 is strongly non-arithmetic in the sense that

$$\limsup_{|\lambda| \to \infty} |\int e^{i\lambda x} dF_0(x)| < 1.$$

Similar results hold if F_0 is arithmetic provided the definitions of ρ_0, $\rho_{1,}$, γ_0, and γ_1 are appropriately modified.

Our first result is a refinement of Theorems 10.13 and 10.16 of Siegmund (1985).

_Proposition 1._Let $\theta_0 < 0 < \theta_1$ and $\Delta_i = \theta_i - \theta_i^*$ as above. Suppose $\theta_i \to 0$ and $b \to \infty$ in such a way that $\Delta_i b$ converges to a finite, non-zero limit. Then

$$(2) \quad E_{\theta_i}(\tau) = [\Delta_i \psi'(\theta_i)]^{-1}\{\exp[-\Delta_i(b + \rho_1 - \rho_0)] + \Delta_i(b + \rho_1 - \rho_0) - 1\} + \gamma_1 - \gamma_0 + o(1),$$

and

$$(3) \qquad\qquad E_0(\tau) = (b + \rho_1 - \rho_0)^2 + \gamma_1 - \gamma_0 + o(1).$$

Uniformly for $b^{-1}x$ bounded away from 0 and 1

$$(4) \qquad \begin{aligned} E_{\theta_i}^x(\tau) = {}&[\Delta_i \psi'(\theta_i)]^{-1}\{\exp[-\Delta_i(b + \rho_1 - \rho_0)] - \exp[-\Delta_i(x - \rho_0)] \\ &+ \Delta_i(b + \rho_1 - x)\} + \gamma_1 + o(1),\end{aligned}$$

and

$$(5) \qquad\qquad E_0^x(\tau) \doteqdot (b - x + \rho_1)(b + x + \rho_1 - 2\rho_0) + \gamma_1 + o(1).$$

Proof. The expansion (3) is obtained by Siegmund (1975); and (2) to one degree less precision is given by Siegmund (1985, Theorem 10.16). The slightly refined form of (2) given here may be proved by following the approach of Siegmund (1985) in conjunction with the arguments of Theorems 1–3 of Siegmund (1979). The details are omitted.

To obtain (4) and (5), note that for all x and θ

$$(6) \qquad\qquad E_\theta^x(\tau) = E_\theta^x(N) + P_\theta^x\{S_N \le 0\}E_\theta^0(\tau),$$

where $N = \inf\{n : S_n \notin [0, b)\}$. According to Siegmund (1979, Theorem 3)

$$(7) \qquad P_{\theta_i}^x\{S_N \le 0\} = \frac{\exp[-\Delta_i(b + \rho_1 - \rho_0)] - \exp[-\Delta(x - \rho_0)]}{\exp[-\Delta_i(b + \rho_1 - \rho_0)] - 1} + o(b^{-2}),$$

and by a similar argument

$$(8) \qquad \begin{aligned} E_{\theta_i}^x(N) = {}&[\psi'(\theta_i)]^{-1}[(1 - p_i)(b + \rho_1 - \rho_0) - x + \rho_0] \\ &+ \gamma_0 p_i + \gamma_1(1 - p_i) + o(1),\end{aligned}$$

where p_i denotes the right hand side of (7). Substitution of (2), (7), and (8) into (6) proves (4); and (5) may be proved similarly.

4

Remark. If we compare (for example) (2) to the approximation given in Theorem 10.16 of Siegmund (1985), the new approximation is often only marginally better numerically than the old one. Indeed, in the case of a normal mean $\gamma_1 - \gamma_0 = 0$ by symmetry and the two approximations do not differ numerically. However, the error $o(1)$ in (2) is substantially smaller than the $o(b)$ of the previous result. Consequently it seems reasonable to conclude that the accuracy observed numerically for the normal case in Siegmund (1985) is not a lucky accident, but can be expected quite generally. For the exponential distribution as discussed in Section 3, $\gamma_1 - \gamma_0 = -17/18$, and inclusion of this term converts a good approximation into an excellent one.

To obtain an approximation for $E_\nu(\tau - \nu \mid \tau > \nu)$ we first proceed heuristically and then more rigorously. Observe that

$$(9) \qquad E_\nu(\tau - \nu \mid \tau > \nu) = \int_{[0,b)} E_{\theta_1}^x(\tau) P_\nu\{S_\nu - \min_{k \leq \nu} S_k \in dx \mid \tau > \nu\}.$$

According to Pollak and Siegmund (1986), as ν and $b \to \infty$

$$P_\nu\{S_\nu - \min_{k \leq \nu} S_k \in dx \mid \tau > \nu\} \to P_{\theta_0}\{\max_{n \geq 0} S_n \in dx\}$$

in the sense of weak convergence. If we use (4) and this limiting relation in (9), we obtain

$$(10) \qquad \begin{aligned} E_\nu(\tau - \nu \mid \tau > \nu) &\cong [\Delta_1 \psi'(\theta_1)]^{-1}\{\Delta_1(b + \rho_1 - E_{\theta_0} M_1) \\ &\quad - \exp(\Delta_1 \rho_0) E_{\theta_0} \exp(-\Delta_1 M_1) + \exp[-\Delta_1(b + \rho_1 - \rho_0)]\} + \gamma_1, \end{aligned}$$

where we have put $M_1 = \max_{n \geq 0} S_n$. According to Siegmund (1979, Theorem 1)

$$(11) \qquad E_{\theta_0}(M_1) = |\Delta_0^{-1}| - \rho_1 + \frac{1}{2}|\Delta_0|\gamma_1 + o(\Delta_0) \quad (\Delta_0 \to 0).$$

By a similar argument, if θ_0 and θ_1 converge to 0 at the same rate

$$(12) \qquad \begin{aligned} E_{\theta_0} \exp(-\Delta_1 M_1) &= (\Delta_1 - \Delta_0)^{-1}[|\Delta_0| + |\Delta_0|\Delta_1 \rho_1 \\ &\quad - \frac{1}{2}\Delta_0^2 \Delta_1 \rho_1^2] + \frac{1}{2}\Delta_1|\Delta_0|(\rho_1^2 - \gamma_1) + o(\Delta_1^2). \end{aligned}$$

Substitution of (11) and (12) into (10) suggests the approximation, for $\theta_1 > 0$

$$(13) \qquad \begin{aligned} E_\nu(\tau - \nu \mid \tau > \nu) &\cong [\Delta_1 \psi'(\theta_1)]^{-1}\{\Delta_1(b + 2\rho_1 + \Delta_0^{-1}) \\ &\quad + \Delta_0 \exp[\Delta_1(\rho_1 + \rho_0)]/(\Delta_1 - \Delta_0) + \exp[-\Delta_1(b + \rho_1 - \rho_0)]\} + \gamma_1; \end{aligned}$$

and a similar calculation with $\theta_1 = 0$ yields in that case

$$(14) \qquad E_\nu(\tau - \nu \mid \tau \geq \nu) \cong (b + \rho_1 - \rho_0)^2 - 2\Delta_0^{-2}\exp[\Delta_0(\rho_1 + \rho_0)] + \gamma_1.$$

In Theorem 1 below we attempt a mathematically rigorous derivation of (13) and (14). We first let ν and $b \to \infty$ holding θ_0 and θ_1 fixed and obtain an approximation for $E_\nu(\tau - \nu \mid \tau > \nu)$ which contains some difficult to evaluate constants. We then let θ_0 and θ_1 converge to 0 (at the same rate) and obtain simple expressions for these constants in terms of ρ_1, ρ_0, and γ_1. In the case $\theta_1 = 0$ this calculation leads to (14). In the case $\theta_1 > 0$, however, the second exponential on the right hand side of (13) does not survive our mathematically rigorous analysis, because it disappears into the $o(1)$ error term when $b \to \infty$. For most values of b and θ_1 the term is very small, and it makes little difference whether it is included in the final approximation or not; but for small values of θ_1 it can lead to a significant improvement. See the numerical examples of Section 3.

Let $M_0 = \min_{n \geq 0} S_n$, $M_1 = \max_{n \geq 0} S_n$.

Theorem 1. Assume $\theta_0 < 0 \leq \theta_1$. Suppose $\nu \to \infty$ and $b \to \infty$. Then for $\theta_1 > 0$

$$E_\nu(\tau - \nu \mid \tau > \nu) = [\psi'(\theta_1)]^{-1}[b + E_{\theta_1}(S_{\tau_+}^2)/2E_{\theta_1}(S_{\tau_+}) - E_{\theta_0}(M_1)$$

$$(15) \qquad + E_{\theta_1}(M_0)\int_{[0,\infty]} P_{\theta_1}^x\{\tau_- < \infty\}P_{\theta_0}(M_1 \in dx)$$

$$+ \int_{[0,\infty)} E_{\theta_1}^x(S_{\tau_-}; \tau_- < \infty)P_{\theta_0}(M_1 \in dx)] + o(1),$$

and for $\theta_1 = 0$

$$E_\nu(\tau - \nu \mid \tau > \nu) = (b + \rho_1 - \rho_0)^2 + \gamma_1 - \gamma_0 - E_{\theta_0}(M_1^2) + \int_{[0,\infty)} E_0^x(S_{\tau_-}^2)P_{\theta_0}(M_1 \in dx)$$

$$(16) \qquad + 2\rho_0[E_{\theta_0}(M_1) - \int_{[0,\infty)} E_0^x(S_{\tau_-})P_{\theta_0}(M_1 \in dx)] + o(1).$$

If in (15), θ_1 and θ_0 converge to 0 in a fixed ratio, the sum of the four terms following b on the right hand side of (15) equals

$$2\rho_1 + \Delta_0^{-1} + \Delta_0/[\Delta_1(\Delta_1 - \Delta_0)] + \Delta_0(\rho_1 + \rho_0)/(\Delta_1 - \Delta_0)$$

$$(17) \qquad + \frac{1}{2}\Delta_0\Delta_1(\rho_0 + \rho_1)^2/(\Delta_1 - \Delta_0) + \gamma_1\theta_1 + o(\Delta_1).$$

Similarly if $\theta_0 \to 0$ _in (16), the sum of the terms following_ $(b + \rho_1 - \rho_0)^2$ _on the right hand side_

of (16) equals

$$(18) \qquad -2\Delta_0^{-2} - 2\Delta_0^{-1}(\rho_0 + \rho_1) - (\rho_0 + \rho_1)^2 + \gamma_1 + o(1)$$

<u>Proof</u>. We consider only (15) and (17). The proofs of (16) and (18) are similar. To prove (15) it suffices by (9) to find suitable approximations for the different terms on the right hand side of (6) and integrate these with respect to the limiting distribution of $P_\nu\{S_\nu - \min_{k\le\nu} S_k \in dx \mid \tau > \nu\}$. According to Pollak and Siegmund (1986), this limiting distribution is the same as the unconditional limiting distribution

$$(19) \qquad \lim_{\nu\to\infty} P_\nu\{S_\nu - \min_{0\le k\le\nu} S_k \in dx\} = P_{\theta_0}\{M_1 \in dx\}.$$

By the definition of τ, Wald's identity, and (for example) Theorem 9.28 of Siegmund (1985),

$$(20) \qquad \begin{aligned} \psi'(\theta_1)E_{\theta_1}(\tau) = E_{\theta_1}(S_\tau) &= E_{\theta_1}(S_\tau - \min_{k\le\tau} S_k) + E_{\theta_1}(\min_{k\le\tau} S_k) \\ &= b + E_{\theta_1}(S_{\tau+})/[2E_{\theta_1}(S_{\tau+})] + E_{\theta_1}(M_0) + o(1). \end{aligned}$$

as $b \to \infty$. Also by Wald's identity and the renewal theorem, as $b \to \infty$ (recall that $N = \inf\{n : S_n \notin [0,b)\}$)

$$(21) \qquad \begin{aligned} \psi'(\theta_1)E_{\theta_1}^x(N) = &\{b + E_{\theta_1}(S_{\tau+}^2)/[2E_{\theta_1}(S_{\tau+})]\}P_{\theta_1}^x\{S_N \ge b\} - x \\ &+ E_{\theta_1}^x(S_N; S_N \le 0) + o(1). \end{aligned}$$

Substituting (20) and (21) into (6) and integrating with respect to (19) gives (15).

Now suppose that θ_0 and θ_1 converge to 0 in some fixed ratio. It follows from Theorem 1 of Siegmund (1979) and its proof that

$$E_{\theta_1}(S_{\tau+}^2)/[2E_{\theta_1}(S_{\tau+})] = \rho_1 + \gamma_1\theta_1 + o(\Delta_1),$$

(11) holds, and similarly

$$E_{\theta_1}(M_0) = -\Delta_1^{-1} - \rho_0 - \frac{1}{2}\Delta_1\gamma_0 + o(\Delta_1).$$

Hence the sum of the four terms following b on the right hand side of (15) equals

$$(22) \qquad \begin{aligned} &2\rho_1 + \Delta_0^{-1} - (\Delta_1^{-1} + \frac{1}{2}\gamma_0\Delta_1)\int_{[0,\infty)} P_{\theta_1}^x\{\tau_- < \infty\}P_{\theta_0}(M_1 \in dx) \\ &+ \int_{[0,\infty)} E_{\theta_1}^x(S_{\tau_-} - \rho_0; \tau_- < \infty)P_{\theta_0}(M_1 \in dx) + \gamma_1(\theta_1 + \frac{1}{2}\Delta_0) + o(\Delta_1). \end{aligned}$$

7

We consider the first integral in (22); the second may be analyzed similarly and is slightly simpler. Let $\rho_0(\theta) = E_\theta(S_{\tau_-}^2)/2E_\theta(S_{\tau_-})$ $(\theta \leq 0)$. An argument along the lines of Theorem 2 of Siegmund (1979), but with somewhat more care to details shows that for some $K > 0$

$$(23) \qquad P_{\theta_1}^x\{\tau_- < \infty\} = \begin{cases} \exp[-\Delta_1(x - \rho_0)] + o(\Delta_1^2) + O(e^{-Kx}) \\ \exp[-\Delta_1(x - \rho_0)][1 + \Delta_1 E_{\theta_1^*}^x(S_{\tau_-} - \rho_0(\theta_1^*)) + O(\Delta_1^2)], \end{cases}$$

where $o(\Delta_1^2)$ and $O(\Delta_1^2)$ are uniform in x and $O(e^{-Kx})$ is uniform in Δ_1 provided Δ_1 is small enough. If we split the range of integration into $[0, (\log \Delta_1^{-1})^2)$ and $[(\log \Delta_1^{-1})^2, +\infty)$ and use the first approximation in (23) for large x and the second for small x, we find that

$$(24) \qquad \begin{aligned} -\Delta_1^{-1} \int_{[0,\infty)} P_{\theta_1}^x(\tau_- < \infty) P_{\theta_0}(M_1 \in dx) &= -\Delta_1^{-1} e^{\rho_0 \Delta_1} E_{\theta_0}[\exp(-\Delta_1 M_1)] \\ &\quad - e^{\rho_0 \Delta_1} \int_{[0,\infty)} \exp(-\Delta_1 x) E_{\theta_1^*}^x[S_{\tau_-} - \rho_0(\theta_1^*)] P_{\theta_0}(M_1 \in dx) + o(\Delta_1). \end{aligned}$$

The method of proof of Theorem 1 of Siegmund (1979) gives (12), and by a similar argument

$$(25) \qquad \begin{aligned} \int_{[0,\infty)} E_{\theta_1}^x(S_{\tau_-} - \rho_0; \tau_- < \infty) P_{\theta_0}(M_1 \in dx) &= \frac{1}{2}\Delta_1|\Delta_0|\gamma_0/(\Delta_1 - \Delta_0) \\ &\quad + \int_{[0,\infty)} \exp(-\Delta_1 x) E_{\theta_1^*}^x[S_{\tau_-} - \rho_0(\theta_1^*)] P_{\theta_0}(M_1 \in dx) + o(\Delta_1). \end{aligned}$$

It is easy to see that the integrals on the right hand sides of (24) and (25) are of order Δ_0, and hence when (24) and (25) are added, they cancel up to an error of order Δ_0^2. (This cancellation is very fortunate; otherwise the final answer would contain some constants which we do not know how to evaluate.) Substitution of (12), (24), and (25) into (22) and simplification yield (17).

3. Numerical Examples.

A particularly simple and important special case of the approximations of Section 2 is the normal case: $dF_0(x) = (2\pi)^{-1/2}\exp(-x^2/2)dx$, $\psi(\theta) = \theta^2/2$, $\Delta_i = 2\theta_i$, and as pointed out above

$$\rho_1 = -\rho_0 \cong .583, \qquad \gamma_0 = \gamma_1 = \frac{1}{4}.$$

It is easy to compare the approximations of Proposition 1 with values determined numerically by Dobben de Bruyn (1968) or Lucas and Crosier (1982) to see that the approximations are very accurate. We have performed a Monte Carlo experiment in order to obtain an indication

8

of the accuracy of the approximations in (13)–(14). The results are given in Table 1. The Monte Carlo experiment had 2500 repetitions, and the results are reported $\pm$ one standard error. Figures given in parentheses are theoretical calculations.

The theoretical approximations appear to be excellent, with the approximation for large ν being very good even for $\nu = 10$.

<div align="center">

Table 1

$E_\nu(\tau - \nu \mid \tau > \nu)$

$b = 4.83,\ |\theta_0| = .5\ (E_{\theta_0}(\tau) \cong 790)$

</div>

$\theta_1 \backslash \nu$	0	10	50	200	$\to \infty$
1.0	$5.6 \pm .04$ (5.5)	$5.1 \pm .04$	$5.2 \pm .05$	$5.1 \pm .05$	(5.1)
.50	$10.1 \pm .1$ (10.0)	$9.2 \pm .1$	$9.3 \pm .1$	$9.3 \pm .1$	(9.2)
.25	$16.1 \pm .2$ (16.4)	$15.6 \pm .2$	$15.2 \pm .2$	$15.9 \pm .2$	(15.3)
.00	$37.0 \pm .6$ (36.0)	$33.8 \pm .6$	$33.4 \pm .6$	$34.2 \pm .6$	(34.2)

<div align="center">

$b = 8.14,\ |\theta_0| = .25,\ (E_{\theta_0}(\tau) \cong 794)$

</div>

	0	10	50	200	$\to \infty$
.50	$16.7 \pm .1$ (16.6)	$15.2 \pm .2$	$14.4 \pm .1$	$14.5 \pm .2$	(14.2)
.25	$29.3 \pm .3$ (29.3)	$27.1 \pm .3$	$25.8 \pm .3$	$26.3 \pm .2$	(25.5)
.12	$45.9 \pm .6$ (45.6)	$42.2 \pm .6$	$42.1 \pm .6$	$41.7 \pm .6$	(40.5)
.00	89.3 ± 1.4 (86.6)	84.8 ± 1.4	80.2 ± 1.4	83.9 ± 1.4	(78.9)

The approximations of Section 2 allow us to make the following simple assessment of Lucas an Crosier's (1982) fast initial response feature. As a result of some numerical experimentation they suggest starting the CUSUM process at $x = b/2$. For the first test in Table 1, for $b = 4.87$ and $x = 2.44$, we have $E_{\theta_0}^x(\tau) \cong 790$. Also, for $\theta_1 = |\theta_0| = .5$, $E_{\theta_1}^x(\tau) \cong 5.9$ and $E_\nu(\tau - \nu \mid \tau > \nu) \cong 9.3$ for large ν. Hence we have obtained about a 40% reduction in the average delay to detect an immediate change at the price of a 1% increase in the average delay to detect a change which occurs later. It follows that the fast initial response feature with $x = b/2$ results in an overall savings if $\nu = 0$ occurs at least $2\frac{1}{2}\%$ of the time. (This conclusion changes slightly if we make the comparison at other values of θ_1.)

A second class of important, simple examples involves detection of a change in the mean of

an exponential distribution, or what is almost the same, a change in the intensity of a Poisson process. A systematic discussion with extensive numerical tables is given by Lucas (1985). For illustrative purposes consider detection of a change of the parameter λ of an exponential distribution ($\lambda =$ mean^{-1}) from λ_0 to $\lambda_0^* > \lambda_0$. To put this problem in the canonical form of Section 2 we set

$$dF_0(x) = \exp(x - 1)dx, \qquad x \leq 1$$
$$0 \qquad\qquad x > 1,$$

$\theta = \lambda/\tilde{\theta} - 1$, and $\psi(\theta) = \theta - \log(1 + \theta)$, where $\tilde{\lambda} = (\lambda_0^* - \lambda_0)/\log(\lambda_0^*/\lambda_0)$. In terms of the exponentially distributed observations, say $y_1, y_2, \cdots$, we have $x_k = 1 - \tilde{\lambda}y_k$. Note that $\theta = 0$ corresponds to $\lambda = \tilde{\lambda}$.

It follows easily from the lack of memory property of the exponential distribution that the P_0-distribution of S_{τ_-} is unit exponential on $(-\infty, 0)$. Hence from the Wiener-Hopf factorization of the distributions of S_{τ_+} and S_{τ_-} (e.g. Siegmund, 1985, Theorem 8.41) one easily obtains that the P_0-distribution of S_{τ_+} is uniform on $(0, 1)$. It follows from simple calculations that $\rho_1 = \frac{1}{3}$, $\gamma_1 = \frac{1}{18}$, $\rho_0 = -1$, and $\gamma_0 = 1$. (If instead of the exponentially distributed y's one observes continuously a Poisson process $N(t)$, so $S_n = \Sigma(1 - \tilde{\lambda}y_k)$ is replaced by $N(t) - \tilde{\lambda}t$ in the definition of τ, then $\theta = \log(\lambda/\tilde{\lambda})$ and by continuity $\rho_0 = \gamma_0 = 0$. If our test is to detect a decrease in λ, then ρ_0 and ρ_1 are interchanged, as are γ_0 and γ_1.)

Table 2 gives a brief comparison of the approximations of Proposition 1 with numerically determined results of Lucas (1985), which are given in parentheses. The quality of the approximations is excellent.

Recently a number of authors (e.g. Woodhall, 1983) have developed algorithms for numerical computation of the entire distribution of τ. Analytic approximations to this distribution will be discussed in a future paper.

10

Table 2

Exponential Distribution: $E_\theta^x(\tau)$

$(\lambda_0 = 1,\ \tilde{\lambda} = 1.25)$

Parenthetical Entries form Lucas (1985)

		λ		
b	x	1.0	1.25	1.5
7	0	295 (295)	68.5 (68.8)	33.8 (33.9)
8	0	474 (471)	86.2 (86.3)	39.6 (39.6)
7	3.5	259 (259)	49.2 (49.5)	20.8 (20.9)
8	4.0	424 (422)	62.2 (62.3)	24.1 (24.1)

REFERENCES

van Dobben de Bruyn, C. S. (1968). *Cumulative Sum Tests*, Griffin, London.

Lucas, J. M. (1985). Counted data CUSUM's, *Technometrics* **27**, 129–144.

Lucas, J. M. and Crosier, R. B. (1982). Fast initial response (FIR) for cumulative sum quality control schemes, *Technometrics* **24**, 199–205.

Pollak, M. (1987). Average run lengths of an optimal method of detecting a change in distribution, *Ann. Statist.* **15**.

Pollak, M. and Siegmund, D. (1985). A diffusion process and its application to detecting a change in the drift of Brownian motion, *Biometrika* **72**, 267–280.

Pollak, M. and Siegmund, D. (1986). Convergence of quasi-stationary to stationary distributions for stochastically monotone Markov process, *Jour. Appl. Probab.*

Siegmund, D. (1975). Error probabilities and average sample number of the sequential probability ratio test, *J. Roy. Statist. Soc. Ser. B* **37**, 394–401.

Siegmund, D. (1979). Corrected diffusion approximations in certain random walk problems, *Adv. Appl. Probab.* **11**, 701–719.

Siegmund, D. (1985). *Sequential Analysis: Tests and Confidence Intervals*, Springer-Verlag, New York-Berlin-Heidelberg-Tokyo.

Woodhall, W. H. (1983). The distribution of the run length of one-sided CUSUM procedures for continuous random variables, *Technometrics* **25**, 295–301.

END

11- 86

DTIC